知识管理论坛 ISSN 2095-5472 CN11-6036/C
Knowledge Management Forum

E-mail: kmf@mail.las.ac.cn
http://www.kmf.ac.cn

【学术探索】

国内科研机构和高校机构知识库建设现状调研与对比分析

张伶[1,2,3] 祝忠明[1] 寇蕾蕾[1,2] 姚晓娜[1]
1. 中国科学院西北生态环境资源研究院 兰州 730000
2. 中国科学院大学经济与管理学院图书情报与档案管理系 北京 100049
3. 新乡医学院图书馆 新乡 453000

摘要：[**目的 / 意义**] 在 IR 发展进入平稳期及向下一代 IR 转型的重要节点，梳理我国 IR 的建设历程、现状和面临的挑战，以达到推动我国 IR 加速和深入发展的目的。[**方法 / 过程**] 首先采用网络调查法及电话、邮件调查法相结合的实证调研法，以国内高校和中国科学院、中国农业科学院等科研院所 IR 为调查对象，对全国 IR 发展现状进行调查梳理；其次对调研结果按照高校系统、科研系统和总体情况进行对比分析，总结 IR 发展成绩与问题。[**结果 / 结论**] IR 建设取得了一定成就，在建及拟建 IR 数量不断增加，资源规模大幅增长。然而 IR 发展存在管理政策不完善、访问限制过多、知识产权不明确、开放性不足、自存储率低等问题。

关键词：机构知识库 科研机构 高校 建设现状
分类号：G250.7

引用格式：张伶 , 祝忠明 , 寇蕾蕾 , 等 . 国内科研机构和高校机构知识库建设现状调研与对比分析 [J/OL]. 知识管理论坛 , 2020, 5(2): 122-134[引用日期]. http://www.kmf.ac.cn/p/207/.

1 引言

自 2002 年《布达佩斯开放获取倡议》正式发布以来，开放获取已取得了可喜的成就。近一年来可以说是开放获取迅速发展的一段时期，随着欧洲开放获取“S 计划”的推出，能否打破付费墙，实现即时开放获取已成为学术界普遍关注的焦点。而机构知识库（Institutional Repository，IR）作为开放获取实现的重要途径之一，在经过十多年的持续增长发展后，其数量和内容都已初具规模，在促进学术成果的开放获取中发挥了积极作用。但就科研人员和学

基金项目：本文系中国科学院文献情报能力建设专项“机构知识库持续运行建设及开放科研知识云”（项目编号：Y8ZG051001）研究成果之一。
作者简介：张伶 (ORCID：0000-0002-0995-0753)，博士研究生；祝忠明 (ORCID：0000-0002-2365-3050)，主任，研究馆员，博士生导师，通讯作者，E-mail: zhuzm@llas.ac.cn；寇蕾蕾 (ORCID：0000-0002-0953-7125)，博士研究生；姚晓娜 (ORCID：0000-0001-6167-4138)，馆员，硕士。
收稿日期：2020-02-29 发表日期：2020-04-21 本文责任编辑：刘远颖

知识管理论坛，2020(2):122–134
DOI：10.13266/j.issn.2095–5472.2020.012

术界对 IR 的认知、认可、支持和使用效果来看，IR 仍未能充分融入科研学术过程和体系，特别是在开放获取实践的力度提速加强、深度和广度不断拓展提升并加速向开放科学迈进之际，IR 面临新的挑战和机遇，处于向下一代 IR 转型的重要节点。

发展 IR 一直以来是实现开放获取的重要形式之一，也是支持开放获取 S 计划落实的重要渠道和手段之一。跟欧美发达国家相比，我国开放获取运动发展滞后而缓慢。尽管我国在 2018 年 12 月举行的柏林开放获取会议上发布了立场声明，明确表示支持 OA2020 和开放获取“S 计划”，但目前尚无具体的推进实施措施。在全球开放获取深入发展、开放科学蓬勃兴起的背景下，对我国 IR 发展现状进行全面调查，梳理我国 IR 的建设历程、现状和面临的挑战，既可以回首来路，认清不足，总结其发展成绩与问题，又可以总结经验，展望其未来发展趋势，以达到推动我国 IR 加速和深入发展的目的。

2 数据来源、研究方法和研究过程

2.1 数据来源

为力求全面、系统、客观地展现我国 IR 建设发展的全貌，了解我国机构知识库整体建设现状，对中国高校系统及科研系统 IR 的发展现状进行调查梳理。笔者收集了来自中国科学院（以下简称“中科院”）、中国农业科学院（以下简称“农科院”）、CHAIR、ROAR、OPENDOAR 以及高校的数据，汇总形成支持统计分析的数据基础。根据调查数据，从基本建设情况、政策制定情况和应用情况等方面，对我国 IR 的建设和发展进行了全方位多维度的梳理和分析。

2.2 研究方法和研究过程

首先采取网络调查法，采集各高校网站及已发表文献中提及到的 IR 信息，其次采用电话调研及邮件调研方式逐个核实该 IR 是否存在。对于互联网可访问的 IR，其数据信息由网络实证调研及电话邮件调研相结合的方式获取。对于互联网不可访问，不对外提供服务的 IR，采取问卷调查、文献检索、电话和邮件调研等多种方式进行 IR 关键数据信息调查、核实和完善。由于中国科学院及中国农业科学院的 IR 都在各自的 IR 联盟系统，同时 IR 联盟维护及时，数据清晰，因此科研机构 IR 数据较高校情况好统计分析。根据调查数据，从总体概况、建设服务状态、资源建设、构建平台功能与服务、开放程度、政策规范、联盟发展等方面，对我国 IR 的建设和发展进行了全方位多维度的梳理和分析。

3 数据分析

3.1 国内科研机构和高校机构知识库基本建设情况

3.1.1 服务状态对比分析

截至 2019 年 11 月，我国共有 IR325 个。其中，高校系统 166 个，科研院所 152 个，其他系统 7 个。从运行服务状态来看，正常运行服务的 IR 共有 287 个，有 18 个正在建设中，20 个已处于服务中止状态。正常运行和服务访问的 IR 中，高校系统 131 个，科研系统 149 个，其他系统 7 个。而真正支持互联网公开访问的 IR 有 209 个，其中，高校系统 92 个，科研系统 110 个，其他系统 7 个。具体服务状态如图 1 所示：

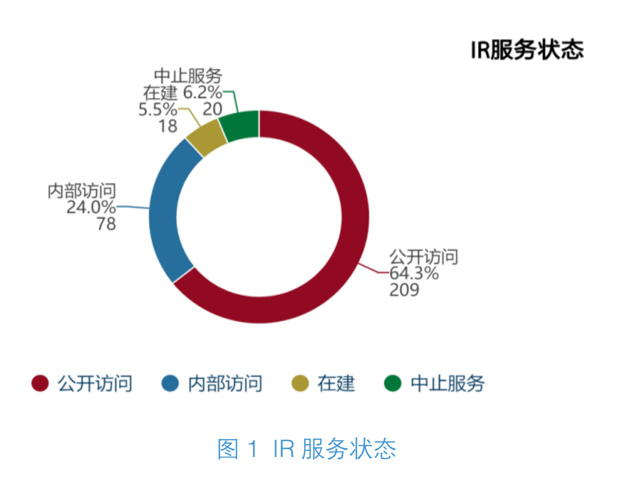


图 1 IR 服务状态



知识管理论坛
2020 年 第 2 期（总第 26 期）

42 所“双一流”大学中共有 IR36 个，其中正常运营的有 26 个，在建 6 个，中止服务的有 4 个。“双一流”高校 IR 建设状态如表 1 所示：

表 1 “双一流”高校 IR 建设服务状态

序号	高校名称	IR 网址	上线时间 / 年	数据规模 / 条	全文量 / 篇	状态
1	厦门大学	http://dspace.xmu.edu.cn	2006	166 828	/	公开访问 全文开放
2	中国农业大学	http://202.112.175.14:4237	2008	100 000	/	公开访问
3	北京大学	http://ir.pku.edu.cn	2010	540 694	301 347	公开访问
4	电子科技大学	http://ir.uestc.edu.cn	2013	20 000	/	校内访问
5	西安交通大学	http://www.ir.xjtu.edu.cn	2013	327 585	/	公开访问
6	中国人民大学	http://ir.lib.ruc.edu.cn	2013	200 846	78 260	公开访问
7	兰州大学	http://ir.lzu.edu.cn	2014	158 610	31 177	公开访问
8	清华大学	http://ir.lib.tsinghua.edu.cn	/		/	内部访问
9	西北工业大学	http://ir.nwpu.edu.cn	2015	180 882	62 990	公开访问
10	北京大学医学部	http://ir.bjmu.edu.cn	2016	148 631	58 551	校内访问
11	北京航空航天大学	http://ir.lib.buaa.edu.cn	2016	180 977	88 572	公开访问
12	北京理工大学	http://bitssp.dayainfo.com	2016	225 004	/	公开访问
13	同济大学	http://ir.tongji.edu.cn	2016	321 947	120 000	校内访问
14	中南大学	http://csu.organ.yunscholar.com	2016	362 085	/	公开访问
15	中山大学	/	2017	480 000	/	校内访问
16	东北大学	/	2018	100 000	/	校内访问
17	湖南大学	http://hnussp.dayainfo.com	2018	251 153	/	公开访问
18	华中科技大学	/	2018	10 000	/	校内访问
19	武汉大学	http://openir.whu.edu.cn	2018	421 647	/	公开访问
20	大连理工大学	http://dlutir.dlut.edu.cn	2019	294 486	/	公开访问
21	东南大学	http://58.192.117.79	2019	300 000	/	校内访问
22	复旦大学上海医学院	http://shmuir.fudan.edu.cn	/	252 241	/	公开访问
23	哈尔滨工业大学	http://hit2ssp.dayainfo.com	/	256 819	/	校内访问
24	山东大学	http://ir.lib.sdu.edu.cn	/	425 659	/	公开访问
25	天津大学	http://tjussp.dayainfo.com	/	360 500	/	公开访问
26	重庆大学	http://cqu.irtree.com	/	244 108	/	公开访问 全文开放
27	上海交通大学	/	/	/	/	在建
28	北京师范大学	/	/	/	/	在建
29	南开大学	/	/	/	/	在建
30	四川大学	/	/	/	/	在建
31	中国海洋大学	/	/	/	/	在建
32	复旦大学	/				在建
33	华东师范大学	/	/	/	/	中止服务
34	浙江大学	/	/	/	/	中止服务
35	南京大学	/	/	/	/	中止服务
36	云南大学	/	/	/	/	中止服务

注：数据获取时间为 2019 年 11 月 20 日

知识管理论坛，2020(2):122-134
DOI：10.13266/j.issn.2095-5472.2020.012

高校系统 IR 服务状态存在这样几种情况：①图书馆网站显示有 IR 的多家院校，其链接无法访问，且通过电话、邮件的方式均联系不到，因此此类机构并未算入 IR 总数。②此次调研发现部分 IR 转入内部访问甚至停止服务的现象。例如，电子科技大学、同济大学现在变为内部访问；华东师范大学、宁夏大学、石河子大学、浙江大学等的 IR 经笔者电话调查确认已经中止服务。③有一批特殊类型的 IR，即运营商自己做了一批 IR 给机构试用，机构实际并没有做，未得到机构的承认。例如，维普曾做过一批类似的 IR，如湖南师范大学 IR、华南师范大学 IR 等。本次调研结果，对这类 IR 做了排除。④部分 IR 长期未更新，如湖南大学 IR、北京建筑大学 IR2019 年以来无数据更新，有些 IR 访问量及数据量都很少。⑤不支持全文公开访问和下载，是高校 IR 建设的普遍现象。⑥支持公开访问的 IR，有不少存在平台不稳定、不能正常打开的情况。

科研系统及其他系统 IR 的运行服务情况相对明了，全部 159 个科研院所 IR 中，中国农业科学院系统和其他系统的所有 IR 均正常运营且可公开访问；中国科学院 114 个 IR 中，目前正常运营且可公开访问的有 72 个，内部访问的有 39 个，中止服务的有 3 个。

3.1.2 上线时间对比分析

我国 IR 的建设起步总体比国外较晚，高校系统和科研系统的发展也有较大差异。2006 年厦门大学建成我国第一个 IR，其后几年时间，高校 IR 的建设进展相对缓慢。2013-2018 年是高校 IR 较快发展的时期，有 80 个 IR 完成建设和上线。与高校系统不同，科研系统 IR 的建设呈现鲜明的批量建设方式和特征。中国科学院是国内科研系统最早启动建设 IR 的机构，从 2007 年开始以中国科学院力学所 IR 和中国科学院国家科学图书馆 IR（现中国科学院文献情报中心 IR）示范建设为起步，2009 年开始规模化推广建设，2013 年基本完成所属大部分研究所 IR 平台的建设上线。中国农业科学院系统的 IR 发展较晚，但建设上线速度很快，从 2017 年开始建设，2018 年即完成了全部农科院系统 IR 的批量建设和上线。依据调研的 IR 建设时间数据，我国 IR 建设发展的时间历程如图 2 所示：

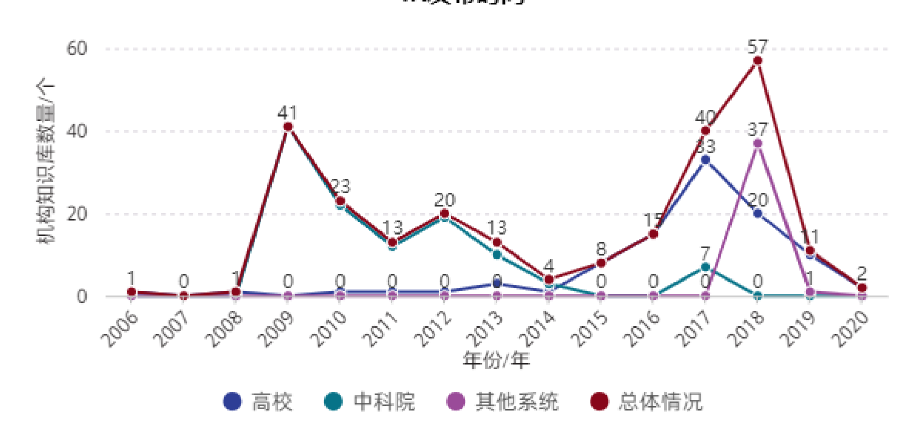


图 2 IR 发布时间分布

3.1.3 地区分布对比分析

所有 IR 的地区分布统计结果如图 3 所示。从中可以看出，中国大陆地区各省市自治区都建设有 IR，但发展分布很不均衡。



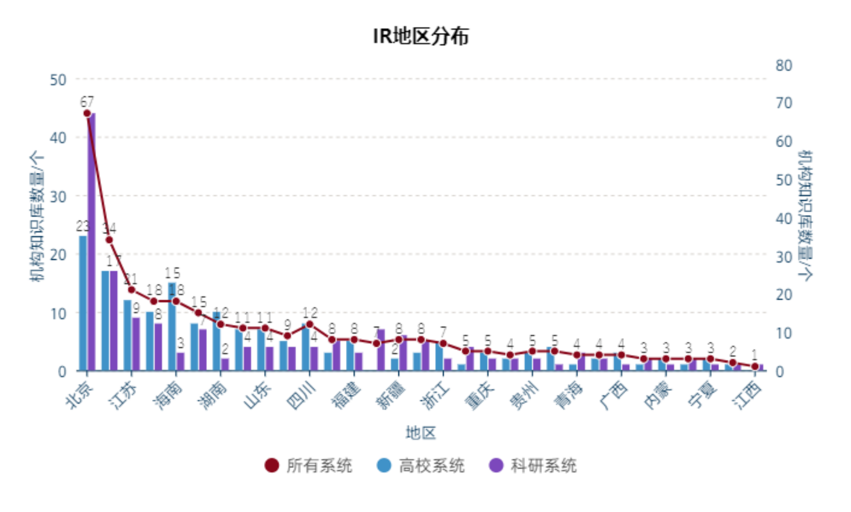


图 3 IR 地区分布

北京、上海和江苏是 IR 数量排行最靠前的 3 个省份，三地 IR 之和占到我国 IR 总数的三分之一强。经济发展水平高的东部地区和两湖地区，高校和科研机构数量也多于其他地区，有 220 个 IR，占全国 IR 总量的 67.7%。相应地，中西部地区高校和科研机构数量相对较少，经济发展水平相对落后，IR 数量也较少，共有 105 个，只占全国 IR 总量的 32.3%。高校系统和科研系统 IR 的地区分布状况与总体的区域分布状况基本相同，区域发展和分布的不均衡一样存在。

3.1.4 资源建设类型对比分析

我国 IR 资源收录内容丰富，收录了多种类型的数据资源，但主要以公开发表的文献为主，灰色文献及非文献资源占比较少。中国科学院机构知识库网格和农业机构知识库联盟统计有资源类型分布，且 94% 的科研院所 IR 都包含在这两个联盟，因此联盟的情况可以代表科研院所的情况；“双一流”高校的情况代表了我国高校的较高水平，因此高校的情况以可访问的“双一流”高校为样本，统计出各类型资源类型占比，如图 4 所示：

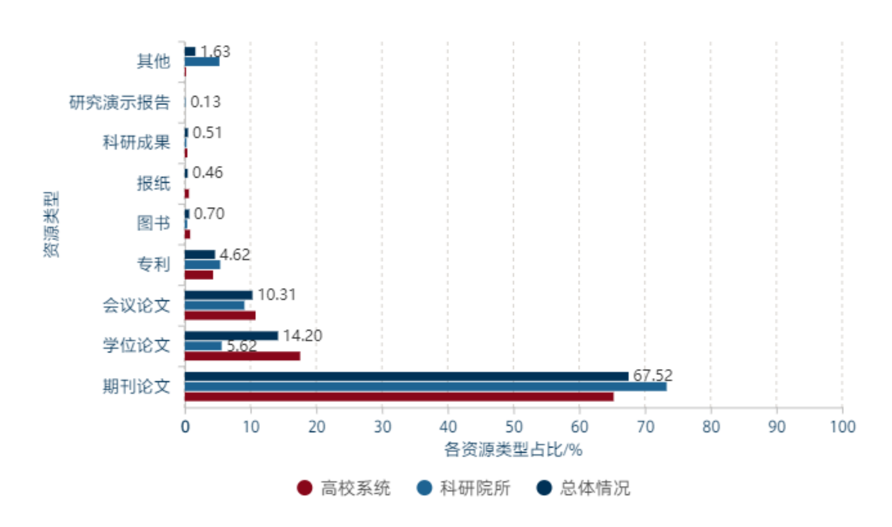


图 4 机构知识库资源收录类型

知识管理论坛，2020(2):122-134
DOI：10.13266/j.issn.2095-5472.2020.012

由图 4 可知，IR 收录资源比例最大的是期刊论文、学位论文和会议论文，这 3 种类型的资源占比达到了 92%。非文本的资源、多媒体资源收录较少，比如录音材料、视频资料和工具软件等较少看到，非正式成果的保存现状堪忧，IR 尚未发挥应有作用。科研院所 IR 对比高校 IR 情况可以看出，科研院所 IR 收录的资源数量较高校少，但类型更多样化，有一些平时不太常见的资源类型，比如考察资料、传媒扫描、大科学装置等。从资源覆盖的学科或主题范围看，高校 IR 的学科和主题范围基本与所属高校的学科和研究范围重合，大多属于多学科范畴。中国科学院的 IR 收录的资源与所属研究所相关，多属于学科 IR 范畴，理科比文科数据量更丰富，开放度也更高一点。

3.1.5 构建平台及提供服务对比情况

（1）平台软件类型。国内 IR 软件平台建设多选择自主开发或第三方软件平台。在可公开访问的 IR 中，以中国科学院和中国农科院为代表的科研系统 IR 全部采用了自主开发的软件平台；高校 IR 中，只有北京大学 1 家采用自主开发的模式，其余 89 家选择了第三方软件平台。第三方软件平台可提供成熟的功能套件，并提供一定的定制开发和数据服务，极大降低了 IR 建设的技术门槛，是越来越多机构建设 IR 软件平台的主要选择。只有厦门大学 IR 和广西民族大学 IR 在采用 DSpace 开源软件。国内目前主要的 IR 第三方软件平台有 CSpace、超星、维普、同舟云、西安知先、湖南维度等，其分布如图 5 所示：

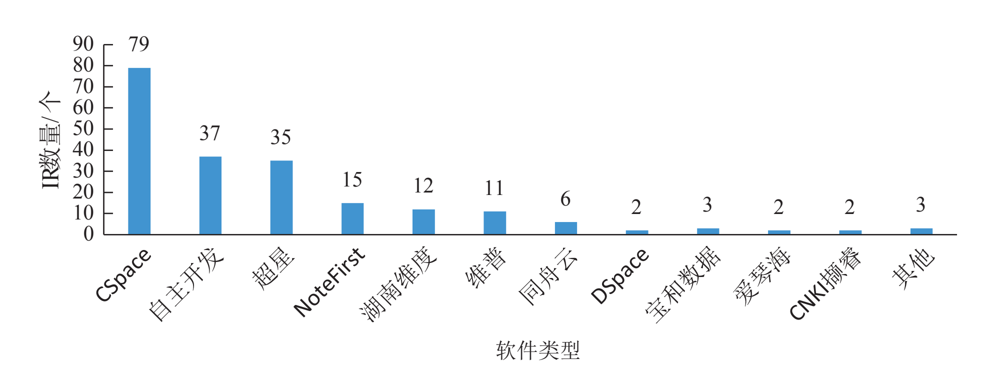


图 5 国内 IR 软件类型分布

（2）平台软件功能服务。尽管国内各 IR 采用的软件平台有差异，但各平台提供的核心功能基本一致，主要包括知识成果的存缴及管理、知识成果的发现利用服务、科研成果数据与相关科研实体数据汇集与开放互操作服务、个人学术主页、统计分析及图谱化揭示等方面。对各 IR 网站逐个调研，并基于软件类型进行归纳概括，每种软件类型 IR 的基本服务功能和扩展服务功能如表 2 所示。

各软件平台均提供最基本的知识成果存缴管理和知识资源发现利用服务，提供多种知识成果的组织及浏览方式。各软件平台提供的 IR 成果基础发现功能大同小异，体现为分类导航、一站式检索等基础功能，引用、推荐、收藏、评注、导出、分析等支持用户利用的服务功能。部分 IR 支持通过与学校相关科研信息系统进行数据共享、交换与关联集成，实现学校科研成果数据与相关科研信息的关联利用和服务。

为吸引研究人员参与，CSpace、超星、维普、西安知先等平台开发个人学术主页，提供包括个人简介、科研成果、成果统计、关联作者、研究热点、可视化分析等服务。同时，CSpace、维普支持导出学者分析报告，同舟云支持通过微信二维码查看学者报告且支持查看 RDA 关联数据格式。为获得领导部门支持，IR 增设学科分析功能，致力于为我国高校学科评

估、学科能力建设提供支持建设，CSpace、维普、超星、同舟云等 IR 建设平台功能体系中均有统计分析及图谱化功能，支持多场景组合条件下的数据分析及支持导出多种格式的报表。

表 2 IR 软件类型的服务功能概览

软件类型	基本服务功能	扩展服务功能	IR示例
CSpace	研究单元&专题、作者、文献类型、学科分类	知识图谱、新闻公告、知识整合、学习讨论、学科评估、Google Scholar 索引	上海科技大学知识管理系统、兰州大学、中国科学院机构知识库服务网格IR
西安知先	成果提交、分类浏览、我校学者、收录情况、统计	学术期刊、成果管理、学科分析、友好期刊、个人科研助理、团队科研协同	上海大学IR、中国人民大学IR、大连理工大学IR
维普	机构、学者、成果、学科、分析、动态概况	团队、学科分析、成果管理、学术分析	成都理工大学IR、成都中医药大学IR、贵州理工学院IR
超星	专家学者、机构、分类浏览、收录情况、统计分析	科研管理系统、学科分析报告	安徽大学IR、东北农业大学、大连海事大学、天津大学
湖南维度	成果展示、学者列表、院系列表	科研分析	华中师范大学、西南交通大学、中南大学
宝和数据	学者、成果、机构、统计分析	团队	山东大学、华东师范大学
同舟云	成果展示、院系分析、机构学者	学科建设、知识图谱、统计报表、顶级期刊	上海中医药大学、复旦大学上海医学院

不同 IR 平台提供不同类型的个性化、细粒度的增值服务和扩展功能，CSpace 扩展功能包括移动设备的兼容适应、新闻公告、学术讨论厅，偏向于用户体验、知识分享及交流服务；西安知先则包括学术期刊推荐、个人科研助理、团队科研协同等服务，偏重于科研管理服务。总体而言，目前我国大部分 IR 服务功能从最初的信息搜集、存储和展示，逐渐转向对服务教学、科研管理和学科建设提供支持，将 IR 转变为进行教学和科研活动的日常场所。

3.2 机构知识库政策制定情况

综合性的政策框架对 IR 的持续建设和运行服务至关重要。国际通行的做法是，在 IR 平台公开发布 IR 政策，供相关各方参考使用。对公开访问的 IR 的调研结果显示，有 32 家机构在平台发布了政策，其中高校 30 家，科研及其他系统 2 家。需要说明的是，IR 平台发布的政策存在模板套用和政策内容完全一致的现象，某种程度上是因为 IR 平台本身提供一种通用政策框架，这些政策是否得到机构本身的认可和成为正式的政策，本次调查中笔者未对这种现象进行进一步核实。总体上，我国制定和公开发布 IR 政策的机构数量不多，管理运行工作有待进一步规范。

3.2.1 高校 IR 政策制定情况

在 90 家正常运营且可公开访问的高校 IR 中，有 30 家在平台（约 33%）提供了政策信息，由于部分高校 IR 使用了相同的政策模板，政策内容完全一致，因此将其归为一类。对这些政策内容进行归纳分析后，发现共有 9 类包含实质性内容的政策，主要涉及内容政策、提交政策、保存政策、全文及元数据使用政策、撤回政策、隐私政策等方面（见表 3）。

目前高校 IR 政策有几个方面的问题：①无院校级政策，政策多为图书馆部门或 IR 出台的管理措施，导致其政策实施力度不够；②数据政策不详不尽，多数内容政策都笼统地指出允许用户以个人学习、研究等非盈利目的，对存储在 IR 中的作品全文进行免费获取和使用，严禁机器恶意批量下载全文，使用过程注明引用信息，任何形式或媒介的全文商业销售必须得到版权所有人的正式许可等；③目前，高校制定激励人员存缴的政策的不多，未发现强制性举措；如何激励学者将学术成果存缴到

知识管理论坛，2020(2):122-134
DOI：10.13266/j.issn.2095-5472.2020.012

IR 中并及时开放共享，是个严峻的挑战；④版权政策不明晰，大部分高校 IR 提交者限定为在校人员，成果可以由作者本人或作者授权的委托代理人在任何时间进行提交，作品版权由原作者或第三方持有，任何版权的侵犯由作者全权负责。

表 3 我国可公开访问高校 IR OA 政策概览

政策内容 / 高校	内容政策	提交政策	保存政策	撤回政策	全文政策	元数据政策	许可政策	隐私政策	免责声明	权利义务	服务政策
北京大学	✔	✔	✔	✔	✔	✔	✔	✔			
华中师范大学等12所高校[1]	✔	✔	✔	✔	✔	✔		✔	✔	✔	
上海大学等8所高校[2]	✔	✔	✔	✔	✔	✔		✔			
西安交通大学 内蒙古大学 山东大学	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
大连理工大学西安工程大学	✔	✔	✔	✔	✔	✔		✔	✔		
兰州大学	✔	✔	✔				✔				
厦门大学	✔	✔	✔		✔	✔					
武汉大学	✔	✔	✔	✔	✔	✔		✔			
海南省高校IR联盟高校[3]	✔	✔	✔	✔	✔	✔	✔	✔	✔		

注：1- 包括华中师范大学、西南交通大学、中南大学、湖南城市学院、湖南工业大学、衡阳师范学院、湖南农业大学、湖南中医药大学、南华大学、武汉轻工大学、武汉理工大学、中南林业科技大学。
2- 包括上海大学、西北工业大学、中国人民大学、河北大学、江苏大学、武汉科学技术大学、中国石油大学（北京）、北京航空航天大学。
3- 包括海南大学、海南师范大学、海南热带海洋学院、海南医学院、琼台师范学院、三亚学院、海南职业技术学院、海南政法职业学院、海南外国语职业学院、海南软件职业学院、海南科技职业学院、海南工商职业学院、三亚航空旅游职业学院（平台未显示政策，但均属于海南省高校 IR 联盟成员）

3.2.2 科研院所 IR 政策制定情况

相比高校相对缺乏明确公开的开放获取政策声明，中国科学院和中国农业科学院都制定了院级的开放获取政策，为 IR 的内容建设提供了有利的制度保障。2014 年 5 月，中国科学院发布了《中国科学院关于公共资助科研项目发表的论文实行开放获取的政策申明》[1]，要求中科院系统的研究人员及研究生在学术期刊上所发表的研究论文，作者应在发表后 12 个月内开放获取。中国农业科学院于 2019 年 7 月也发布了《中国农业科学院关于公共资助科研项目发表的论文实行开放获取政策（草案）》，并做出了类似的声明。中国科学院和中国农业科学院也都依据国家《科学数据管理办法》分别出台了各自支持科学数据开放共享的实施办法[2-3]，为 IR 下一步介入和参与科学数据的开放共享提供政策参考依据。

在具体的 IR 实施政策方面，中国科学院 IR 在建设过程中出台了一系列支持研究所 IR 建设应用的政策体系：推出系列 IR 建设制度政策规范参考指南，《IR 建设管理办法》《公共教育科研单位 IR 内容存缴与传播的权益管理政策指南》等。除院级 IR 政策之外，科研院所 IR 提供有详实的人员激励政策，如中科院 IR 的建设情况纳入全院年度科研信息化评估指标，部分研究所将其作为年度科研成果统计奖励基础和依据，研究所机构网站人才队伍栏目依托 IR 个人学术主页功能建设，利用 IR 知识资产统计服务来支持有关科研成果的统计分析和辅助支持有关评估应用工作；中国农业科学院的 IR 也借助各院所科研处，将科研管理、项目申报、成果评价一体化，将 IR 发展与科研绩效评价挂钩。

这些激励政策在一定程度上促进了 IR 的发展和研究人员积极性。

3.3 国内科研机构和高校机构知识库应用现状

3.3.1 数据规模

由于 IR 内部访问和资源统计信息不公开现象比较普遍，调查反馈的数据存在模糊性表述的情况也比较多，结合各公开访问 IR 的数据，粗略统计，我国 IR 收集各类科研成果信息的总量在 1400 万条，主要为公开发表论文，具体统计结果如表 4 所示：

表 4 我国 IR 收录数量规模统计

元数据总量	高校系统IR		科研院所IR	
	数量/个	比例/%	数量/个	比例/%
≥10万	44	40.7	2	1.8
<10万，≥1万	46	42.3	56	35.2
<1万且≥5000	8	7.4	44	27.7
<5000且≥1000	6	5.6	39	24.5
<1000	4	3.7	18	11.3
合计	108	100	159	100

所有可公开访问的 IR 中，资源量排在前 10 位的见表 5。由于高校的规模和体量的关系，可以看出排在前面的 IR 都属于高校系统。

表 5 IR 资源量 TOP10

排名	机构	资源总量/条
1	北京大学	540 694
2	山东大学	425 659
3	武汉大学	421 647
4	中南大学	370 829
5	天津大学	360 500
6	西安交通大学	327 585
7	大连理工大学	298 765
8	东华大学	261 333
9	复旦大学医学院	252 241
10	湖南大学	251 153

3.3.2 全文收录情况

可统计的全文科研成果 200 多万件，占到总资源量的 14.4%，可开放获取全文科研成果近 30 万件，仅占总资源量的 2% 多一点。因部分 IR 全文量统计获取问题，实际全文科研成果和开放获取全文成果预计会略高于目前的统计结果。根据 IR 全文资源量的规模，开放访问且可获取全文量的 IR 可以归为如下 5 个级别，具体分布结果统计如表 6 所示：

表 6 IR 全文资源规模分布统计

全文资源总量	高校系统IR		科研院所IR	
	数量/个	比例/%	数量/个	比例/%
≥10万	2	8.70	0	0
<10万，≥1万	9	39.13	20	26.67
<1万且≥5000	3	13.04	18	24.00
<5000且≥1000	3	13.04	30	40.00
<1000	6	26.09	7	9.33
合计	23	100	75	100

对于正常建设和持续运行的 IR，全文收录比和全文开放比更能反映 IR 对全文收录和服务能力。IR 全文收录比排名靠前的机构如表 7 所示：

表 7 IR 全文收录比 TOP10

序号	机构	资源量/条	全文量/篇	全文收录比/%
1	上海微系统与信息技术研究所	13 171	13 171	100.00
2	上海应用物理研究所	11 247	11 247	100.00
3	动物研究所	6 780	6 780	100.00
4	昆明动物研究所	7 581	7 568	99.83
5	广西民族大学	9 433	9 405	99.70
6	科技战略咨询研究院	4 086	4 067	99.53
7	寒区旱区环境与工程研究所	12 745	12 638	99.16
8	国家授时中心	2 318	2 284	98.53
9	兰州化学物理研究所	10 883	10 712	98.43
10	上海高等研究院	195	191	97.95

注：IR 全文收录比 =IR 全文量 /IR 资源量

知识管理论坛，2020(2):122–134
DOI：10.13266/j.issn.2095–5472.2020.012

可以看出高校系统 IR 在资源总量上占优势，但在资源的开放获取服务方面不如科研院所。高校 IR 的全文收录率较低，大部分不到 50%，且相当一部分高校 IR 仅提供论文成果的原文链接，甚至有些连全文链接也没有。直接提供全文下载的高校 IR 仅有厦门大学、重庆大学、广西民族大学等少数几家。两大 IR 建设体量较大的科研系统中，农业 IR 联盟的所有 IR 目前未收集全文，中国科学院系统的 IR 在全文收录和开放获取服务方面进展较好，部分研究所的全文收录率和全文开放率可达 100%，大部分研究所的全文收录率都可达 50% 以上。

3.3.3 开放程度

IR 的开放性和国际化程度对于 IR 促进机构成果的开放获取和开放传播极为重要，主要从 IR 收录全文的开放度、元数据开放收割性、国际化登记和收录性、跨语言界面和环境等方面对我国 IR 的开放性和国际化程度进行分析。

（1）互联网可访问性。除去已经中止服务和在建的 IR，剩余的全部 287 个正常运行和服务的 IR 中，支持互联网公开访问的 IR 共有 209 个，占 72.8%；不支持互联网公开访问即仅支持机构内部访问的 IR 共有 78 个，占 27.2%。在 131 个正常运行和服务的高校 IR 中，支持互联网公开访问的有 92 个，仅支持校内访问的有 39 个。在 156 个正常运行和服务的科研系统 IR 中，支持互联网公开访问的有 117 个，仅支持内部访问的有 39 个。

（2）元数据开放收割性。OAI（Open Archives Initiative）协议，是一种独立于应用的、能够提高 Web 上资源共享范围和能力的互操作协议标准。目前，可公开访问的 IR 中，支持 OAI-PMH 接口的共有 74 家，占比 36%，其中高校仅有两家，分别为北京大学和上海科技大学，科研院所共有 73 家且大部分为中科院系统所属机构，仅有 1 家其他机构（南京水利科学研究院）。

（3）全文开放度。在正常运行服务的 IR 中，支持全文开放获取的 IR 仅有 38 个，占比 13.2%。公开访问的高校 IR，大部分功能都有 IP 限制，需注册用户才可正常访问，多数高校 IR 未统计全文收录情况，不提供全文开放下载，仅部分 IR 提供有全文 DOI 号或全文在外网的链接。只有中科院 IR 全部统计有全文数量和可供开放的全文数量，按照 IR 开放获取全文量/IR 资源量的方式计算 IR 全文开放比，有 3 个 IR 达到 100% 的全文开放率，占到中科院 IR 总数的 2.6%，有 23 个 IR 提供的全文开放率达到 60% 以上，占到中科院 IR 总数的 20.2%。对中科院 IR 全文访问状态进行统计，共有暂不开放、限制开放、开放获取、延迟开放 4 种状态 [4]，如图 6 所示：

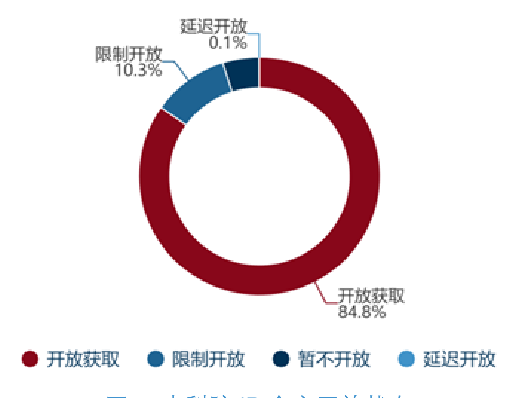


图 6 中科院 IR 全文开放状态

进一步统计限制全文开放的情况，发现不同类型资源开放度不同，针对不同文献类型，数据访问状态同样不同，可提供全文开放获取的一般都是公开发表的期刊论文等，学位论文、工作文档等多数仅提供元数据信息，不提供全文。如自然科学史研究所 IR 对期刊论文提供开放获取，对学位论文则仅提供元数据信息。

4 讨论与分析

4.1 IR 联盟管理形式具有优势

通过科研机构与高校系统的对比可以发现，科研系统由于均处于 IR 联盟体系中，机构之间通过政策、规范、技术、人力资源等要素的共建共享，可以集约化的方式促进 IR 的快速建设。

以中国高校 IR 联盟、中国科学院 IR 网格和农业 IR 联盟为代表的 IR 联盟，在组织和推动我国高校及科研系统 IR 的建设发展过程中，通过适宜的合作机制建设、标准规范共建和技术资源共享，已经发挥了显著作用。IR 联盟作为一种 IR 联合服务的模式，通过成员 IR 之间的互操作和数据共享，汇聚构建联合服务的 IR 网络，充分发挥知识集聚效应和网络效应，极大地提升 IR 的推广服务效果和应用影响。

中国高校机构知识库联盟 CHAIR，通过标准规范共享、免费提供软件平台和技术支持服务，对高校 IR 的发展起了很大的推动作用，引起了高校 IR 建设的热潮，但相对于科研院所 IR 联盟的发展还可以更进一步。首先，CHAIR 联盟的参建机构数量与高校 IR 规模相比，体量较小，需要积极扩大覆盖面；其次，也存在成员 IR 的正常运行和开放访问率低的问题。CHAIR 联盟平台目前登记的高校 IR 有 50 个 [5]，但正常运营的有 30 个，占联盟 IR 的 60%；支持开放访问的只有 19 个；内部访问不对外提供服务有 11 个，占到了正常运营库的 42.8%。

在全球 IR 逐步进入转折发展的关键时期，我国 IR 联盟亟需更新理念，创新模式，以继续更好地推动我国 IR 面向开放科学环境的升级发展。目前我国的 IR 联盟基本上都采用分布式模式构建，这种模式保证了每个 IR 成员的较高自主权，各成员机构可以自行建立、运营和维护自身的 IR，展现本机构的研究特色与研究能力。然而，较为单一的分布式 IR 联盟也存在一定的缺点：①每个成员机构参与成本相对较高，需要专门投入资金、人力、技术等资本因素，从而不利于中小规模的学术 IR 成员的成长；②成员 IR 具有不一致性，无论从软件的选择还是元数据标准应用及数据收录标准上，各成员机构均会出现差异性。下一步，高校和科研系统的主要联盟之间应进行有力度的合作，在争取国家政策支持和扶植发展方面，形成一致行动并共同努力；在开放获取和开放科研文化的宣传推广方面，进行系统的规划，持续开展活动。

4.2 完整的管理政策和完善的服务体系是保障

高校 IR 的数据管理政策多采用软件提供的统一模版，由图书馆发布，且大部分 IR 未见公开可行的校级 IR 政策，使 IR 从项目任务转为常规任务后的可持续发展存在很大的不确定性。大部分的 IR 内容政策与版权政策不明晰，作者不明确自己是否有权提交最终版本，作者可提供何种权限给予读者，机构不知自己是否有权对外开放获取。同时我国大部分 IR 都由图书馆牵头建设，多数归于技术部负责，还有一些归由资源建设部负责，很少有 IR 指定专人负责管理与运营。但是 IR 的运营与服务本身具有很高的技术含量，这就导致多数 IR 不能保证提供 7×24 小时的稳定访问，死链、空链、错链的情况非常常见。

完善的 IR 政策代表着来自机构层面对实现机构知识资产管理和服务的支持和承诺，是界定 IR 与机构相关应用服务边界和关联嵌入关系的依据，也是 IR 实现机构知识管理需求、定义流程及开展日常管理和运作服务的政策支撑。要争取国家层面的政策支持，资助机构出台强制性措施，科研学术机构出台相应激励措施，3 个方面形成共振。同时应梳理、理清国家对著作权的规定和文件，借鉴国际通用的 CC 许可使用，借鉴国内外高校对 IR 版权处理的相关规定，搜集与整理国内外不同出版社、编辑部、数据库商的版权政策、自存储政策，制定本机构的切实可行的管理政策和服务体系。

4.3 扩大资源存储类型，加强机构知识库资源建设

调研结果显示，数据建设成为我国 IR 建设发展的瓶颈问题，甚至成为风险、失败因素。虽然 IR 的数量和资源建设总量在增加，但是资源建设情况呈现两极分化现象，科研和高校系统 IR 普遍存在资源可用性低，且主要以回溯正式出版的期刊论文、会议论文为主，资源深度和广度有待加强，整体 IR 全文可获取率都不高，限制全文获取是我国 IR 服务的常态。

IR 要使自身发展贴合用户需求，就要加强 IR 的内容资源建设。IR 学术资源包括正式

知识管理论坛，2020(2):122–134
DOI：10.13266/j.issn.2095–5472.2020.012

出版的文本类知识资源和未出版的灰色文献及非文本的多媒体知识资源和隐性知识，IR 的价值和意义更多地通过后者实现。IR 资源建设的新方向是在集成外部资源（outside-in）的同时重点建设推广内部资源（inside-out），基于中转推送 OA 出版物数据建设服务，比如 JISC Publications Router，中科院 OA 论文推送转发服务系统 iSwitch，对机构公开发表成果、学位论文进行自动监测、采集。

要利用资源和平台优势将 IR 打造成学科服务平台，为用户提供个性化、专业化的知识服务。可从以下两个方面入手：①顺应科研数据管理的需求，支持科学数据管理的政策与流程嵌入、组织模式以及与文献的关联服务，实现 IR 与重大科学装置的成果关联，服务装置、文献、数据、人员、项目等科研实体信息的关联共享与组织服务。②将可管理科研知识对象、内容类型扩展，增加对科学数据、图形图像、音视频、三维对象等非文本科研内容管理，支持数据与相关实体的关联，研究成果的跟踪采集与更新服务，项目产出成果关联报告服务，支持科学数据及相关知识的语义整合管理。目前已有 IR 融入科研管理全过程的实践，将 IR 嵌入学科服务，为研究人员提供更加精准化的知识服务。

4.4 取消不必要的访问限制，加大开放获取实施力度

我国高校 IR 大多提供有限制条件的访问，多数只对校内 IP 范围内的用户提供全文下载权限，更注重实现学校知识资产集中管理与保存的功能，不注重知识的扩散与共享。我国 IR 提供多种访问级别：①可对外提供无限制的全文下载，如大部分中科院 IR；②可提供部分资源的无限制全文下载，如山东大学 IR 和中科院部分研究所；③可直接浏览其资源，但不提供下载服务；④仅提供资源全文链接，不提供全文浏览和下载；⑤完全不对外提供服务。不必要的访问限制在促进机构学术成果的可见性和影响力提升方面难以发挥应有作用，对融入国际化发展和应用不够重视，降低了 IR 的可用性及全球影响力，大部分 IR 系统成为实实在在的本地应用。

应取消不必要的访问限制，对不同类型的资源根据其资源属性及作者需求提供相应的开放措施，如期刊论文、会议论文、学术专著及章节等公开或半公开发表的学术成果，可参考著作权法或通过取得作者授权等方式合理地放开全文限制访问；而学位论文、培训课件及内部资料等灰色文献，很多内容属于暂时不公开发表的成果，可以采用元数据公开、全文限制获取方式使用。

5 结语

本文对我国大陆地区科研机构和高校 IR 的总体情况进行了调研，分析了我国 IR 的基本建设情况、政策制定情况和 IR 应用情况，分析当前 IR 发展可改进的方面，并提出相应地解决方案，以期为 IR 发展提供参考。总体来说，我国 IR 的建设取得了一定成就，2019 年我国 IR 整体规模持续增长，在建及拟建 IR 数量不断增加，中国 IR 的总量已位列世界前 10，资源规模大幅增长，服务效果和效应逐步提升。然而，我国 IR 发展还有很大潜力，绝大部分高校还没有建设 IR，42 所“双一流”高校中只有 17 所建设了 IR，从整体上来说，现有 IR 的资源建设水平和服务功能都有待完善。IR 发展存在管理政策不完善、访问限制过多、知识产权不明确、开放性不足、自存储率低等一直未能有效解决的老问题；也涉及在国际知识库发展至新阶段之后面临的新挑战和新问题，如多元复杂开放科学内容的采集管理、主动融入与动态互操作、机器智能内容服务等。

IR 发展到现阶段，其所承载的职能已不再仅是管理和传播该机构的数字资源，更应具有更深层的作用，是促进知识成果开放共享、争夺学术交流话语权、提供技术及科研支撑的重要组成部分。IR 在发展过程中，需要借鉴供给侧思维来调整功能和服务建设的目标，提升有效知识、服务供给的能力和水平，成为有用的



知识管理论坛
2020 年 第 2 期（总第 26 期）

资源系统和技术平台，淘汰“落后产能”，废止无效应用、服务供给，创造新供给，引领新需求，提供新服务。应建立完善的管理与服务体系，加强内容资源建设，以科研人员需求为核心及出发点。进一步借鉴企业数据中台的理念，使 IR 成为机构的学术数据中台，将 IR 建设融入科研人员科研过程，成为以科研人员为核心的科研信息、知识集成管理中心和科研人员日常工作流的一部分。展望未来，IR 需借力开放获取运动和开放科学加速发展的东风，从面向机构智能开放知识基础设施、支持数字学术的开放获取平台环境、融入和支持开放科学过程等多方面提升应用服务价值及用户体验，在促进与支持数字学术和开放科学发展中发挥不可替代的作用。

参考文献：

[1] 中国科学院 . 中国科学院关于公共资助科研项目发表的论文实行开放获取的政策申明 [EB/OL].[2019-12-23].http://www.cas.cn/xw/yxdt/201405/P020140516559414259606.pdf.

[2] 中国农业科学院 . 中国农业科学院开放获取与科学数据管理办法印发 [EB/OL].[2019-12-23].http://aii.caas.cn/bsdt/zhdt/216208.htm.

[3] 中国科学院 . 中国科学院科学数据管理与开放共享办法（ 试 行 ）[EB/OL].[2019-12-23].http://www.cas.cn/tz/201902/P020190220358041915907.pdf.

[4] 中国科学院 . 中国科学院机构知识库网格 .[EB/OL].[2019-12-23]. http://www.irgrid.ac.cn/report?type=summary.

[5] 中国高校机构知识库联盟 [EB/OL].[2019-12-23].http://chair.calis.edu.cn/.

作者贡献说明：

张 伶：负责论文的数据获取、提纲与撰写；

祝忠明：负责论文的修订；

寇蕾蕾：负责文字修订；

姚晓娜：负责数据调研。

Investigation and Comparative Analysis of Institutional Repository in Chinese Scientific Research Institutions and Universities

Zhang Ling[1,2,3] Zhu Zhongming[1] Kou leilei[1,2] Yao Xiaona[1]

1. Northwest Institute of Eco-Environment and Resources,Chinese Academy of Sciences, Lanzhou 730000

2. Department of Library Intelligence and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100049

3. Xinxiang Medical University Library, Xinxiang 453000

Abstract: [Purpose/significance] At the important node of IR development entering a stable period and in the transition to the next generation of IR, this paper reviews the construction process, status quo and challenges of institutional repository in China, so as to achieve the purpose of accelerating and deepening the development of IR in China. **[Method/process]** First, it adopts the empirical research method combining the network survey method and telephone and email survey methods, and takes domestic universities, Chinese Academy of Sciences, Chinese Academy of Agricultural Sciences and other scientific research institutes as the survey object to investigate the development status of the national IR; second, the survey results According to the comparative analysis of the university system, scientific research system and the overall situation, summarize the IR development achievements and problems. **[Result/conclusion]** Some achievements have been made in IR construction. The number of IR under construction and to be built is increasing, and the resource scale is increasing greatly. However, there are some problems in the development of IR, such as imperfect management policies, excessive access restrictions, unclear intellectual property rights, insufficient openness and low self storage rate.

Keywords: Institutional Repository scientific research institutions universities construction status